



GPS différentiel basé vision : Amélioration de la fusion SLAM/GPS à l'aide des modèles 3D des bâtiments

Dorra Larnaout, Vincent Gay-Bellile, Steve Bourgeois, Michel Dhome

► To cite this version:

Dorra Larnaout, Vincent Gay-Bellile, Steve Bourgeois, Michel Dhome. GPS différentiel basé vision : Amélioration de la fusion SLAM/GPS à l'aide des modèles 3D des bâtiments. *Reconnaissance de Formes et Intelligence Artificielle (RFIA)* 2014, Jun 2014, France. hal-00989150

HAL Id: hal-00989150

<https://hal.science/hal-00989150>

Submitted on 9 May 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GPS différentiel basé vision : Amélioration de la fusion SLAM/GPS à l'aide des modèles 3D des bâtiments

D. Larnaout¹ V. Gay-Bellile¹ S. Bourgeois¹ M. Dhome²

¹ CEA, LIST, LVIC, Point Courrier 173, F-91191, Gif-Sur-Yvette, France

² Institut Pascal, UMR 6602 Université Blaise Pascal/CNRS/IFMA

¹ prénom.nom@cea.fr, ² michel.dhome@univ-bpclermont.fr

Résumé

Nous proposons dans cet article une solution originale pour la localisation d'un véhicule en zone urbaine. Celle-ci exploite uniquement des matériels peu coûteux : une caméra VGA, un GPS standard et des données issues d'un Système d'Information Géographique (SIG) tel que celui de l'IGN ou d'OpenStreetMap. Comme d'autres auparavant, notre approche repose sur un module de fusion d'un SLAM visuel avec les données GPS. Cependant, en plus de l'estimation du mouvement du véhicule, notre solution se distingue par l'estimation et la correction du biais affectant les données GPS. A la manière d'un GPS différentiel, ce biais est estimé dynamiquement en comparant la reconstruction SLAM avec les modèles 3D de ville. Les données GPS, ainsi corrigées, sont alors utilisées par le module de fusion. Il en résulte une précision de localisation jusqu'ici inaccessible avec un GPS standard. Ceci a permis son exploitation dans des applications de Réalité Augmentée.

Mots Clef

SLAM, GPS, SIG, milieu urbain dense.

Abstract

We propose in this paper to fuse a vision-based SLAM with GPS data, digital elevation model (DEM) and 3D building models to obtain an accurate 6 degrees of freedom vehicle localization in dense urban area. The fusion proceeds in two steps. First, vision-based SLAM is combined with GPS data and DEM to prevent scale factor drifts. Then, the resulting 3D points cloud is compared with the buildings models to locally estimate and correct the uncertainty of the GPS. These corrected GPS data are thereafter used in the first fusion module to improve the localization accuracy.

Keywords

SLAM, GPS, GIS, dense urban environment.

1 Introduction

Pour se localiser en ville, la majorité des solutions commercialisées se base sur les systèmes GPS. Cependant dans les milieux urbains, la précision de ces capteurs se dégrade

à cause des canyons urbains¹. Pour pallier ce problème, les solutions basées sur un SLAM visuel [1] semblent une alternative prometteuse. En plus de l'estimation des six degrés de liberté de la caméra mobile, il fournit une carte 3D de la scène observée. Toutefois, la localisation assurée par le SLAM visuel n'est pas géo-référencée et présente souvent des dérives. Pour remédier à ces limitations certaines approches fusionnent le SLAM visuel avec les mesures GPS, par exemple [5]. Malgré l'amélioration apportée par la fusion, ces solutions offrent globalement une précision de localisation *dans le plan*² sensiblement équivalente à celle du GPS et donc insuffisante en milieux urbains denses.

Afin d'améliorer d'avantage la précision de cette fusion, nous proposons d'utiliser des données GPS plus précises. Pour y parvenir sans utiliser un capteur coûteux (eg. GPS RTK), nous proposons d'estimer et de corriger dynamiquement le biais affectant les données GPS en comparant la reconstruction SLAM issue de la fusion avec les modèles 3D des bâtiments. Le principe de notre approche est donc semblable à celui du GPS différentiel où les antennes géo-référencées sont remplacées par les modèles 3D des bâtiments. Les données GPS, ainsi corrigées, sont alors utilisées dans le module de fusion pour une meilleure précision de localisation.

Après un bref état de l'art des méthodes de localisation basée vision dans la section 2, nous introduisons le principe général de notre méthode de localisation basée sur deux modules de fusion dans la section 3. Le principe du module de correction des données GPS est détaillé dans la section 4. Des évaluations à la fois sur des données synthétiques et d'autres réelles sont présentées dans la section 5.

Notations. Dans cet article, nous nous basons sur les travaux de [8] qui supposent que les incertitudes du GPS peuvent être modélisées par un biais $\mathbf{b} = (b_x \ b_y)^T$ local auquel est additionné un bruit gaussien de plus faible amplitude $\mathbf{v}_g = (v_{gx} \ v_{gy})^T$. Le SLAM visuel utilisé [1] est basé sur le principe des images clé. La reconstruction 3D résultante contient alors M points 3D $\{\mathcal{Q}_i\}_{i=1}^M$ et N poses de

1. Réflexion du signal GPS sur les façades de bâtiments

2. Le déplacement dans le plan de la route de la caméra et son angle lacet

caméra. La j^{me} pose de caméra est définie par son orientation \mathcal{R}_j et sa position 3D \mathbf{t}_j dans le repère monde. Les modèles des bâtiments représentent un ensemble de H plans $\{\Pi^h\}_{h=1}^H$ correspondant aux façades des bâtiments. Enfin, l'accent $\tilde{\mathbf{x}}$ indique que le vecteur \mathbf{x} est exprimé en coordonnées homogènes.

2 Travaux Connexes

Les solutions existantes de localisation basée sur la vision peuvent être classifiées selon les informations utilisées pour géo-localiser et éviter les dérives de la reconstruction estimée.

La première famille d'approches est basée sur un modèle d'apparence de la scène. Ces approches fonctionnent en deux étapes. La première étape consiste à construire hors ligne une base d'amers géo-référencés précise [2]. Ensuite, un algorithme de reconnaissance de point de vue est utilisé pour assurer la localisation en ligne [3]. La principale limitation de cette famille d'approche est sa sensibilité face aux changements de point de vue et aux conditions d'illumination.

La seconde famille d'approches, par exemple [5] et [9], se base sur la fusion en ligne des contraintes multi-vues du SLAM visuel avec les données GPS. Grâce à cette fusion, ces solutions permettent d'avoir une localisation géo-référencée. Cependant, elles ont deux principales limitations : une précision *dans le plan* limitée à celle du GPS et donc insuffisante en milieu urbain et des fortes imprécisions sur les degrés de liberté restant que nous nommons degrés de liberté *hors plan*. Pour améliorer la précision des paramètres *hors plan*, nous avons proposé dans [7] de fusionner le SLAM visuel avec les données GPS et les contraintes fournies par le MET (*Modèle d'Élévation de Terrain*). Cette solution est basée sur un ajustement de faisceaux incluant une contrainte d'inégalité introduite dans [5]. Ainsi, la fusion n'est prise en compte que si les données GPS et la contrainte fournie par le MET ne dégradent pas significativement les contraintes multi-vues. Malgré l'amélioration de l'estimation des paramètres *hors plan*, la précision de la localisation *dans le plan* reste limitée aux imprécisions du GPS. Pour pallier ce problème, des récents travaux [10] essaient de corriger le biais du GPS en se basant sur les marquages au sol (passage piéton, lignes blanches...). Cependant, en ville cette solution manque de robustesse puisque ces informations sont souvent masquées par d'autres voitures.

Pour éviter les imprécisions liées aux incertitudes du GPS dans les milieux urbains, la troisième famille d'approches propose de se passer totalement du GPS et de se baser plutôt sur des modèles géométriques de la scène. Ces modèles, fournies par les modèles SIG (*Systèmes d'Informations Géographiques*), sont géo-référencés, largement répandus et peuvent être gratuits. Dans [6], nous avons proposé un ajustement de faisceaux contraint à la fois au MET et les modèles 3D des bâtiments. Intégrer ces contraintes directement dans le processus d'optimisation du SLAM

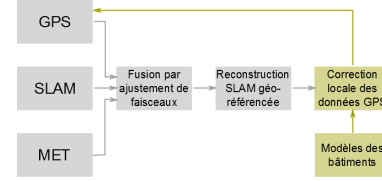


FIGURE 1 – Principe général de l'approche proposée.

En gris la fusion du SLAM visuel avec les données GPS et le MET introduit dans [7]. En vert, le module de correction des données GPS proposé dans cet article afin d'améliorer la précision de la localisation.

a permis d'améliorer considérablement la précision de la localisation. Cependant, cette approche manque de robustesse quand peu de contraintes géométriques sont disponibles (*i.e.* en ligne droite ou quand peu de bâtiments sont observés). De plus, elle demeure sensible à l'initialisation et aux incertitudes des modèles SIG.

Nous proposons dans cet article de tirer avantages des deux dernières familles d'approches pour une localisation précise et robuste en milieu urbain dense. Pour cela un schéma de fusion en deux étapes est proposé combinant un SLAM visuel avec des données GPS, un MET et des modèles de bâtiments.

3 Principe général

Dans cet article, nous nous baserons sur la géo-localisation obtenue à travers la fusion du SLAM visuel avec les données GPS et le MET que nous avons introduit dans [7]. Bien que cette approche assure une certaine robustesse face aux données aberrantes du GPS grâce à la contrainte d'inégalité utilisée dans l'ajustement de faisceaux, elle ne permet pas de pallier le problème du biais (voir figure 6(b)). Contrairement aux données aberrantes, une incertitude sous forme d'un biais local n'entraîne pas une augmentation significative de l'erreur de re-projection au cours de l'optimisation. Par conséquent, la fusion avec les données GPS biaisées est prise en compte, d'où le manque de précision au niveau des paramètres *dans le plan* principalement en milieux urbains. L'utilisation d'un GPS plus précis, tel qu'un GPS différentiel permettrait d'améliorer grandement la qualité de la localisation. Afin d'assurer cette amélioration, ce type de GPS utilise un réseau d'antennes géo-référencées qui transmet au récepteur GPS l'écart entre les positions indiquées par les satellites et leurs positions réelles connues.

Toutefois, le GPS différentiel a des inconvénients. En effet, cette technologie dépend de la présence des antennes géo-référencées qui ne sont pas disponibles partout. Par ailleurs, à cause de son coût très élevé, elle ne représente pas un produit grand public que les utilisateurs des véhicules peuvent exploiter. Enfin, la correction estimée par cette technologie reste sujet à plusieurs erreurs dues notamment au décalage d'horloge et aux réflexions du signal sur les façades des bâtiments.

Dans la suite, nous nous inspirons du principe du GPS différentiel afin de corriger le biais du GPS en exploitant les modèles de bâtiments et ainsi améliorer la précision de la fusion du SLAM avec le GPS et le MET. En d'autres termes, l'écart entre le nuage de points 3D résultant de cette fusion et les modèles des bâtiments est utilisé pour déduire la correction à appliquer localement aux données GPS. Ainsi, dans notre approche, les antennes géoréférencées nécessaires aux GPS différentiels sont remplacées par les modèles des bâtiments. Le schéma général de notre approche est présenté dans la figure 1. Dans la suite nous ne décrivons que le module de correction du biais du GPS. Les détails sur la fusion du SLAM avec le GPS et le MET par ajustement de faisceaux contraint sont dans [7].

4 GPS différentiel basé sur les modèles 3D des bâtiments

Comme il est mentionné ci-dessus, nous cherchons à estimer la correction à appliquer aux données GPS afin d'avoir un nuage de points, issue de la fusion du SLAM avec GPS et MET, aligné localement avec les modèles des bâtiments. Pour ce faire, deux étapes sont exigées :

- **Prédiction de la reconstruction SLAM optimale.** Afin de corriger les erreurs des données GPS, la reconstruction SLAM (poses de la camera et nuage de points), que nous souhaitons obtenir après la fusion par l'ajustement de faisceaux contraint, est prédite.
- **Correction du biais du GPS.** A partir de la prédiction précédente, la correction *dans le plan* à appliquer localement aux données GPS est estimée afin d'atteindre après fusion la reconstruction SLAM optimale.

4.1 Prédiction de la reconstruction SLAM optimale

Afin de corriger les imprécisions du GPS, la première étape de notre processus cherche à estimer la transformation rigide *dans le plan* T_1 qu'il faudrait appliquer localement sur la reconstruction SLAM pour que cette dernière soit plus cohérente avec les modèles 3D des bâtiments, comme le schématise la figure 2. La reconstruction SLAM résultante, appelée reconstruction prédite, n'est pas conservée mais sert uniquement à estimer localement la correction à appliquer aux données GPS (voir section 4.2).

Étant donné que le biais du GPS est local, la transformation recherchée n'est valable que pour un sous ensemble du nuage de points reconstruit. Par conséquent, au lieu de l'estimer à partir de la totalité des points 3D, nous choisissons d'exploiter uniquement les points qui ont été observés dans les n dernières images clé. Un nombre n élevé implique plus de contraintes. Toutefois, ceci peut influencer la précision de la correction en augmentant le risque de modéliser plusieurs biais par un seul. Inversement si n est petit, la correction peut être mal estimée vu le manque de contraintes disponibles. Par ailleurs, pour garantir plus de précision, seuls les points correspondant aux façades des bâtiments

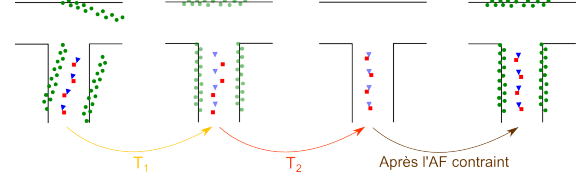


FIGURE 2 – **Schéma global de la solution proposée.** Les points verts représentent le nuage de points correspondants aux façades des bâtiments. Les poses de la caméra sont représentées par les triangles bleus. La reconstruction SLAM optimale prédite est représentée en transparent. Les positions réelles sont opaques. Les positions GPS sont modélisées par les carrés rouges. L'image de gauche représente la reconstruction initiale obtenue par le SLAM contraint au GPS et au MET [7], les poses de la camera ne suivent pas exactement les données GPS. T_1 est la transformation rigide permettant d'obtenir la reconstruction optimale prédite. T_2 représente la correction à appliquer aux données GPS pour tendre après la fusion vers la reconstruction optimale. L'image de droite schématise le résultat obtenu après l'ajustement de faisceaux contraint aux données GPS ainsi corrigées et au MET.

sont utilisés au cours de cette étape³. Nous noterons l'ensemble de points correspondant aux modèles des bâtiments et observés dans les n dernières images clé $\{\mathcal{Q}_i\}_{i \in \mathcal{M}}$. Une fois cet ensemble déterminé, la transformation rigide *dans le plan* T_1 peut être estimée. Pour ceci, nous nous plaçons dans un nouveau référentiel schématisé dans la figure 3. La transformation T_1 dans ce référentiel est obtenue en minimisant la distance euclidienne séparant chaque point de l'ensemble $\{\mathcal{Q}_i\}_{i \in \mathcal{M}}$ à sa façade correspondante. Ainsi, la transformation $T_1 = (\mathcal{R}_1^\Psi | \mathbf{u}_1)$ est obtenue en minimisant l'équation suivante :

$$\mathcal{T}_1 = \sum_{i \in \mathcal{M}} \rho \left(d \left(\Pi^{h_i}, S^{-1} \tilde{\mathcal{Q}}'_i \right) \right), \quad (1)$$

$$\text{où} \quad \tilde{\mathcal{Q}}'_i = \mathcal{R}_1^\Psi S \tilde{\mathcal{Q}}_i + \mathbf{u}_1, \quad (2)$$

avec h_i l'indice de la façade associée au point \mathcal{Q}_i et $d(\Pi, \tilde{\mathcal{Q}})$ est la distance euclidienne séparant le point $\tilde{\mathcal{Q}}$ de

3. Cet ensemble de points est déterminé par la technique de lancer de rayon

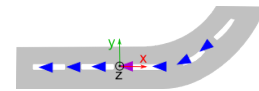


FIGURE 3 – **Repère utilisé pour le calcul de la transformation T_1 .** Les différentes poses de la caméra sont représentées par les triangles bleus. Le triangle violet représente la pose de la caméra définissant l'origine du repère considéré. L'origine du repère est le barycentre des n positions 3D de la camera. Les axes X et Y définissent le plan de la route et l'axe des Z est sa normale.

sa façade II. S est la matrice de passage (3×4) du repère monde au nouveau référentiel. Le M-estimateur Geman-McClure, dont le seuil de rejet est s_1 ⁴, est utilisé pour garantir plus de robustesse face aux données aberrantes.

Par ailleurs, pour garantir plus de robustesse et de précision à notre processus de correction, nous déterminons le nombre de degrés de liberté de la transformation T_1 selon les contraintes disponibles. En effet, comme l'explique la figure 4, si les points 3D considérés correspondent uniformément à des façades orthogonales et latérales par rapport au point de vue de la caméra alors T_1 aura trois degrés de liberté : un pour la rotation \mathcal{R}_1^Ψ i.e. l'angle lacet et deux degrés de liberté pour la translation *dans le plan* $\mathbf{u}_1 = (u_1^x, u_1^y, 0)^T$. Dans le cas contraire, T_1 n'aura que deux degrés de liberté : l'angle lacet et un seul paramètre pour la translation soit u_x soit u_y .

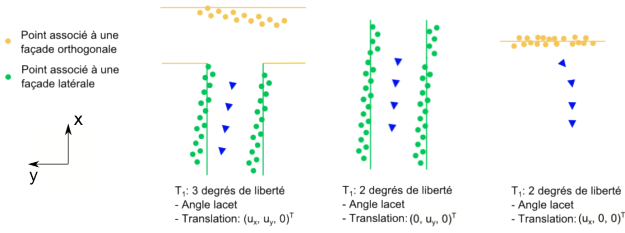


FIGURE 4 – Le nombre de degrés de liberté de T_1 en fonction des contraintes disponibles. Dans la figure à gauche, les points 3D sont uniformément associés aux façades latérales et celle orthogonale par rapport à la direction de la caméra. Dans ce cas, la transformation T_1 possèdera trois degrés de liberté. Dans les deux autres cas, seuls deux degrés de liberté sont estimés.

4.2 Estimation du biais du GPS

La transformation T_1 décrite ci-dessus a permis de prédire la reconstruction optimale. Toutefois, T_1 ne peut pas être appliquée directement aux données GPS puisque les imprécisions de la fusion du SLAM visuel avec le GPS ne sont pas totalement équivalentes de celles du GPS. Ceci est dû au fait que les positions de la caméra $\{\mathbf{t}_j\}_{j=1..N}$ ne suivent pas exactement les données GPS. Par conséquent, à présent, nous cherchons à estimer la transformation rigide *dans le plan* qui, appliquée aux données GPS brutes, permet, à l'issue de l'ajustement de faisceaux contraint, d'avoir les positions de la caméra de la reconstruction SLAM optimale prédite. Une telle transformation T_2 est obtenue en minimisant les distances euclidiennes séparant les données GPS et les positions prédites *dans le plan* de la caméra obtenues en leur appliquant la transformation T_1 comme le montre la figure 2.

Pour calculer la correction recherchée, le même référentiel ainsi que le même nombre de poses de caméra considérées pour le calcul de T_1 sont utilisés. Par ailleurs, contrairement à la transformation T_1 dont le nombre de degrés

de liberté dépend des contraintes disponibles, les trois degrés de liberté *dans le plan* sont estimés pour la correction $T_2 = (\mathcal{R}_2^\Psi | \mathbf{u}_2)$. Ainsi, la fonction de coût résultante est donc donnée par :

$$\mathcal{T}_2 = \sum_{j=N-n}^{N-1} \rho(\|\mathcal{R}_1^\Psi \mathbf{S} \tilde{\mathbf{t}}_j + \mathbf{u}_1 - \mathcal{R}_2^\Psi \mathbf{S} \tilde{\mathbf{t}}_j + \mathbf{u}_2\|, s_2), \quad (3)$$

avec $\tilde{\mathbf{t}}_j = (x_j^{gps}, y_j^{gps}, 0, 1)^T$, x_j^{gps} et y_j^{gps} étant les données GPS correspondant à la j^{me} image clé. De la même façon que précédemment, le M-estimateur Geman-McClure, dont le seuil de rejet est s_2 , est utilisé pour garantir plus de robustesse face aux données aberrantes.

La correction des n précédentes données GPS permet d'améliorer la précision de la localisation obtenue à un instant t , c'est-à-dire une fois que la fusion avec les données corrigées est effectuée. Afin d'éviter les erreurs de localisation qui peuvent se produire à l'instant $t + 1$, la nouvelle donnée GPS doit également être corrigée avant d'être utilisée dans l'ajustement de faisceaux contraint. La connaissance a priori d'un éventuel changement de biais permettrait de prédire avec plus de précision la correction à appliquer aux prochaines données GPS. Toutefois, vu que nous ne disposons pas de cette information supplémentaire pour nos expérimentations (voir section 5), nous choisissons de conserver et d'appliquer la même transformation T_2 préalablement estimée. Ainsi, si aucun changement n'est notable au niveau du biais, cette solution permet de corriger efficacement la prochaine mesure du GPS. Dans le cas contraire, un délai de quelques images clé est nécessaire pour ré-estimer la correction. Malgré ce retard, nous montrerons dans la section 5 que les mauvaises prédictions locales ont peu d'influence sur la localisation finale.

5 Résultats

Pour évaluer la localisation obtenue par notre processus, nous utilisons à la fois des données synthétiques et d'autres réelles. La séquence de synthèse représente un parcours d'un véhicule en milieu urbain. Pour simuler les mesures GPS, les positions *dans le plan* de la caméra, fournies par la vérité terrain, sont perturbées en ajoutant un biais constant par morceaux d'amplitude maximale de 3m auquel est additionné un bruit Gaussien d'amplitude 0.5m. La direction de ce biais change après chaque virage. En ce qui concerne les données réelles, nous utilisons une séquence enregistrée dans le quartier de Versailles (2400m). Pour ceci, le véhicule a été équipé par un GPS standard 1Hz et une simple caméra VGA. Les modèles SIG utilisés ont une incertitude de 2m. Pour chacune des séquences utilisées, nous comparons la précision des localisations obtenues par notre méthode et celles obtenues avec [7]. Dans nos expérimentations, nous fixons le paramètre n à 10. En effet, expérimentalement, nous avons remarqué que cette valeur assure un bon compromis entre précision et robustesse.

5.1 Évaluation sur les données de synthèse

Pour établir notre évaluation, nous comparons dans la figures 5(b) les positions *dans le plan* des données GPS avec

4. Calculer à partir du MAD *Median Absolute Deviation* du vecteur des résidus.

et sans la correction du biais par rapport à la vérité terrain. Nous mesurons également dans la figure 5(c) le résultat de la localisation après la fusion du SLAM et les données GPS avec et sans la correction.

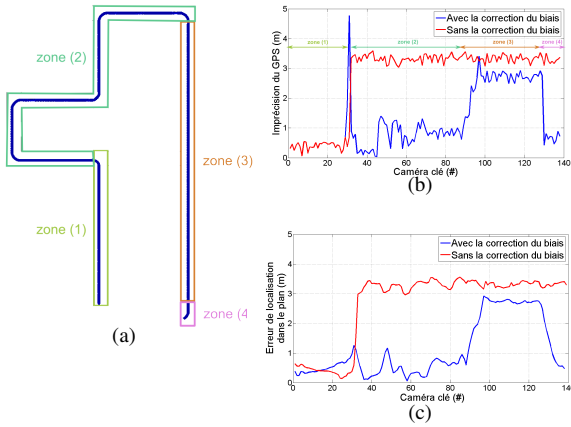


FIGURE 5 – Évaluation du processus de correction du biais du GPS. (a) La trajectoire réelle de la caméra et les différentes zones parcourues. (b) Évolution de l'imprécision des données GPS avec (en bleu) et sans (en rouge) correction. (c) l'évolution de l'erreur de la localisation *dans le plan* après la fusion avec (en bleu) et sans (en rouge) correction.

Comme le montre la figure 5(b), notre méthode permet de réduire notablement les imprécisions du GPS dont la valeur médiane passe de 3.2812m à 0.8527m après correction. Nous remarquons également que le résultat de la correction est fortement dépendant des contraintes disponibles, la direction et l'amplitude de la perturbation. En effet, dans la zone (1) des figures 5(a) et 5(b), seul un bruit gaussien d'amplitude 0.5m est appliqué. Dans ce cas, la contrainte d'inégalité permet de pallier ces imprécisions. Le nuage de points reconstruit est donc aligné avec les modèles des bâtiments et aucune correction n'est estimée.

Quand le biais passe de 0m à 3m, une première correction est effectuée. Ceci permet de réduire considérablement les imprécisions du GPS, comme nous pouvons l'observer dans la zone (2) des figures 5(a) et 5(b). Cette zone est caractérisée par la présence d'importantes contraintes liées aux nombreux virages traversés. Ainsi, la correction est estimée avec précision.

Contrairement à la zone (2), la zone (3) représente une longue ligne droite. Les contraintes disponibles sont alors insuffisantes pour estimer avec précision la correction à appliquer surtout que la direction du biais correspond principalement à l'axe de la route. Par conséquent, les données GPS corrigées sont moins précises que celles de la zone (2). Ces imprécisions sont par la suite corrigées dans la zone (4) grâce aux contraintes supplémentaires apportées par la présence d'un dernier virage.

Malgré cette amélioration globale au niveau de la précision des données GPS corrigées, quelques imprécisions locales sont observées par exemple le pic de la courbe bleue

entre la zone (1) et (2) (figure 5(b)). En effet, vu que la même correction est appliquée sur les n dernières données GPS, certaines mesures n'ayant pas le même biais peuvent être perturbées. Cependant même si le principe de fenêtre glissante de taille n peut causer des imprécisions locales, ces dernières, étant considérées comme des données aberrantes, sont filtrées lors de la fusion grâce à la contrainte d'inégalité de l'ajustement de faisceaux adopté, comme l'illustre la figure 5(c). En effet l'erreur de localisation *dans le plan* à l'issue de la fusion a une évolution similaire à celles des données GPS à l'exception des pics locaux qui disparaissent. Ceci met en évidence l'importance de notre processus de correction qui permet une localisation plus précise. Notons par exemple que l'erreur médiane passe de 3.2635m à 0.6673m en utilisant les données GPS corrigées.

5.2 Évaluation sur les données réelles

Pour évaluer qualitativement la précision de la localisation *dans le plan* pour la séquence réelle, la vue de dessus de la reconstruction obtenue par le SLAM contraint aux données GPS corrigées et au MET est comparée avec celle obtenue en exploitant les données GPS brutes. Par ailleurs, la précision de l'estimation des six degrés de liberté de la caméra est mise en évidence en re-projetant les modèles 3D des bâtiments sur des images extraites de la séquence de Versailles. Concernant l'évaluation quantitative, puisqu'aucune vérité terrain n'est disponible pour la séquence utilisée, nous labellisons manuellement les coins des bâtiments dans quelques images extraites du flux vidéo. Nous calculons par la suite l'erreur de re-projection entre les coins labellisés et la re-projection des coins des modèles des bâtiments.

En comparant les figures 6(a) et 6(b), nous observons que le processus de correction du biais du GPS permet d'améliorer notablement la précision de la localisation *dans le plan*. En effet, après la correction des données GPS, le nuage de points reconstruit est aligné avec les empreintes des modèles des bâtiments. Ceci est également mis en évidence dans la figure 7 où les re-projections des modèles des bâtiments sont mieux recalées sur les façades des bâtiments après la correction. Les résultats quantitatifs du tableau 1 confirment ses observations. En effet, la solution proposée a permis de réduire considérablement la moyenne des erreurs de re-projections qui passe de 14.97 pixels à 8.92 pixels. L'écart-type des erreurs mesurées a également baissé d'une façon notable en utilisant notre méthode de 11.6 pixels à 4.94 pixels. Ceci confirme que l'approche proposée réduit l'incertitude du GPS en milieux urbains garantissant ainsi plus de précision à la localisation après fusion. Cette précision a permis une application de Réalité Augmentée (voir matériel supplémentaire).

6 Conclusion et perspectives

Dans cet article, nous avons introduit le concept de GPS différentiel basé sur les modèles de bâtiments. Notre approche consiste à corriger localement le biais du GPS en comparant le nuage de points 3D issu de la fusion du

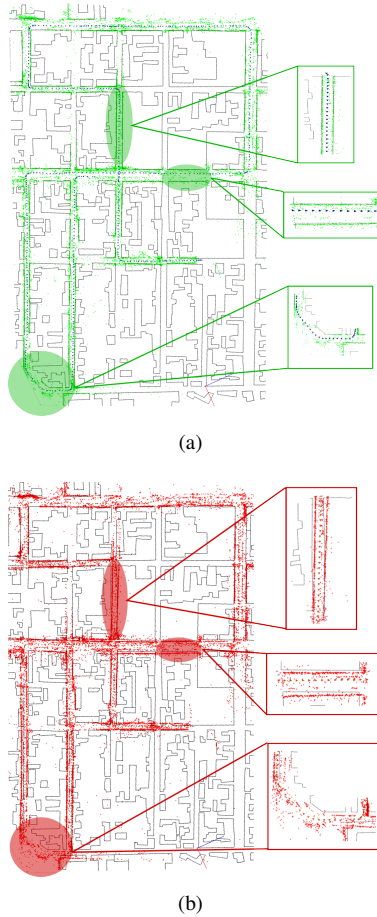


FIGURE 6 – **Vue de dessus des localisations obtenues avec et sans corrections du biais du GPS.** (a) Résultat du SLAM contraint aux données GPS corrigées et au MET. (b) Résultat du SLAM contraint aux données GPS brutes et au MET.

	Erreur de re-projection (pixels)			
	Moyenne	Écart type	Max	Min
[7]	14.97	11.60	44.51	4.16
Solution proposée	8.92	4.94	18.35	4.01

TABLE 1 – **Erreurs de re-projection entre les coins labellisés et les coins des bâtiments re-projetés.**

SLAM visuel et les données GPS avec les modèles de bâtiments. Les données GPS ainsi corrigées ont permis d'améliorer la précision de la localisation par fusion SLAM/GPS dans les milieux urbains denses. Malgré les résultats prometteurs, certaines imprécisions liées au retard de l'estimation de la correction sont notables. Ces limitations seront l'objet de nos prochains travaux. Pour ceci, nous étudierons les changements des constellations des satellites qui peuvent nous informer des changements du biais du GPS comme l'affirme l'étude proposée dans [8]. L'analyse de ces informations permettrait d'estimer avec plus de précision la correction à appliquer et quelles sont les mesures GPS qui sont caractérisées par le même biais.

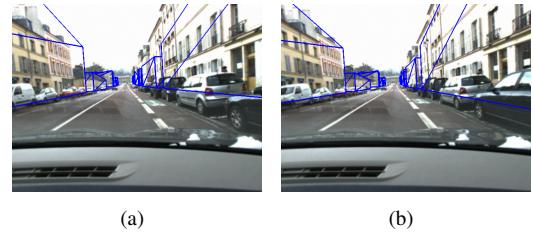


FIGURE 7 – **Exemples de re-projection des modèles des bâtiments.** (a), résultat du SLAM contraint aux données GPS brutes et au MET [7]. (b), résultat obtenu en ajoutant à [7] notre module de correction du GPS. Les erreurs de re-projection associées : (a) 24.10 pixels et (b) 6.15 pixels.

Références

- [1] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser et P. Sayd, *Real Time Localization and 3D Reconstruction*, CVPR 2006.
- [2] Y. Li, N. Snavely, D. Huttenlocher et P. Fua, *Worldwide pose estimation using 3d point clouds*, ECCV 2012.
- [3] A. R. Zamir et M. Shah *Accurate Image Localization Based on Google Maps Street View*, ECCV 2010.
- [4] P. Lothe, S. Bourgeois, E. Royer, M. Dhome et S. N. Collette, *Real-Time Vehicle Global Localisation with a Single Camera in Dense Urban Areas : Exploitation of Coarse 3D City Models*, CVPR 2010.
- [5] M. Lhuillier *Incremental Fusion of Structure-from-Motion and GPS Using Constrained Bundle Adjustments*, PAMI 2012.
- [6] D. Larnaout, S. Bourgeois, V. Gay-Bellile et M. Dhome, *Towards Bundle Adjustment with GIS Constraints for Online Geo-localization of a Vehicle In Urban Center*, 3DIMPVT 2012.
- [7] D. Larnaout, V. Gay-Bellile, S. Bourgeois et M. Dhome, *Vehicle 6-DoF Localization Based on SLAM Constrained by GPS and Digital Elevation Model Information*, ICIP 2013.
- [8] F. Chausse, J. Laneurit et R. Chapuis, *Localisation d'un véhicule sur une carte routière précise*, Traitement de signal 2005.
- [9] D. Schleicher, L. Bergasa, M. Ocana, R. Barea et E. Lopez, *Real-Time Hierarchical GPS Aided Visual SLAM on Urban Environments*, Computer Aided Systems Theory 2009.
- [10] J. Kichun, C. Keounyup et S. Myoungho, *GPS-bias correction for precise localization of autonomous vehicles*, IV, 2013.